

# New Jersey

## Language Arts Literacy

### Curriculum Framework



#### Chapter 6

---

#### Assessment





## ASSESSMENT

As human beings, we are constantly assessing ourselves and others. It is just part of human behavior, perhaps best illustrated by former New York City Mayor Ed Koch, who while in office would ask anyone who would listen, “How am I doing?” Teachers use their professional training and experience to engage in these minute-by-minute assessments in the classroom. Behaviors, such as deciding to ask one student a low-level thinking question and another student a question requiring speculation, may be grounded in our judgment that the first student is capable of only recall while the second is able to engage in higher-order thinking skills. Similarly, the amount of wait time that we allow a student to answer a question may be rooted in our assessment of the student’s ability to think long enough to form an answer. Research has shown in fact that teachers tend to interrupt and correct the miscues of lower-level readers more frequently than they do with their higher-level readers (Goodman, 1991). Here again, the reason seems to be an assessment informing teacher behavior. Often these assessments are unconscious ones, and the teacher is unaware of the impact of the judgment upon learners and their behaviors.

If we are unaware of our minute-by-minute assessments and their impact, perhaps traditional educational practice has contributed to this view by dismissing the importance of the teacher in conventional assessment practices. Traditional practitioners characterized assessment as a periodic measurement based on norm-referenced objective testing (multiple-choice, true-false, and short-answer items) in a single setting; assessment was meant to be teacher-proof. That is, the teacher had no influence over the shape of the test, administration procedures, or the scoring of student responses. This view of assessment has changed dramatically in the last two decades. Roger Farr (1991, p. 490) describes assessment as “best thought of as gathering of a variety of information at diverse times and under differing conditions.” This new view of assessment changes its dimensions. Educational assessment is ongoing and cumulative. It uses open-ended formats. It occurs in a variety of settings. It is theory-based and mainly teacher-mediated.

### Approaches to Assessment

Assessment can be broadly divided into two areas, formal and informal, but as Farr (1991, p. 496) cautions, they really are on a continuum because both are based on student performance. Traditional formal assessment looks at **what** students know at the end of a given period of instruction. Informal assessment looks at **how** a student knows as well as **what** he knows. Formal assessments are usually published. Informal ones are usually teacher-developed although there are published measures, including informal reading inventories, checklists, surveys, and interview guides. Obviously, the measure that we as educators choose determines the information that the instrument will yield. Therefore, we must be very clear about our purpose when we choose an assessment instrument.

The choice of assessment instrument—from teacher observation to student survey to formal published test—should be informed by the assessor’s purpose. Selection of the wrong instrument will not allow inferences appropriate to the assessor’s needs. Traditionally, administrators, seeking information about students’ success in reading, selected published, standardized tests with available normative information, such as the Iowas. This allowed them to compare district performance with statewide and national scores and to comply with Title I requirements. Although the comparisons may have given them confidence in the success of local curriculums, the scores yielded little information that would help guide instruction or curriculum design. With recent changes in programs such as Title I, however, the second purpose for using norm-referenced testing no longer exists.

## Observations

On the other hand, teachers need to derive ongoing data they can use to make decisions about their instructional practices and individual student needs. Educators who work from a constructivist's theory of learning see their students as actively transforming information and experience into personal knowledge. Therefore, they look to sources of information, such as their observations of student behaviors in language arts literacy events, which include individual and small-group activities and conferences as well as students' oral and written products. Teacher observations and written notes serve a variety of purposes, focusing at times on **what** a student has learned and at other times on **how** a student has learned. The purpose for observing informs the method used and the content of the written record. If a teacher's purpose is to assess reader engagement with the text, the teacher might simply observe and note the silent reading behaviors. A teacher may observe a child's invented spelling to see the writer's phonics knowledge and then make note that a lesson on long vowels is appropriate at this point in the learner's progress. A secondary teacher might observe the lack of sentence variety in many student essays and record that finding as a reminder to develop a whole-class lesson on sentence combining. Information derived from teacher observation of the teacher herself, the students, and the context for learning provides potent evidence of student progress in developing literacy. These observations are ongoing and cumulative, and they signal patterns of literacy growth as well as highlight "zones of proximal distance" or areas for future student growth (Vygotsky, 1962, 1978). Teachers use this information to structure opportunities for future learning.

Most important, teachers are at the center of the action: observing day-by-day actions of their students' mood, demeanor, and reactions to school guidelines as well as literacy behaviors. At times, the observation and records will emerge through discussion with the student in a conference. At other times, the focal point of the teacher's observations will be the literacy event. For example, when small groups are discussing literature, teachers may be interested in looking at the dynamics of the group: the degree to which they collaborate and the amount of students' participation as speakers and listeners. In this way, teachers accumulate over time one part of a permanent record documenting students' development. They can use these insights reflectively for purposes of continual curriculum planning and as information to share with students, their parents, and the community. These teacher observations offer far more than the snapshots of a single attainment of a single skill at a single point in time that are provided by standardized tests. Only through teacher observations do we find out how one speaker, listener, writer, reader, or viewer integrates knowledge of discrete skills into complex literate behavior.

## Student Self-Assessment

Another source of information about the student comes from student self-assessment. Hiebert (1991, p. 516) says that traditional classrooms offer too little opportunity for students to participate in setting goals and in monitoring learning. This kind of exercise promotes the students' metacognitive awareness, that is, the language user's awareness of the cognitive demands of literacy, how well the user is comprehending or composing meaning, and what strategies the user might employ to improve language comprehension or composition. There are a number of ways that teachers can promote students' metacognitive awareness through self-assessment experiences. One of these is sitting with students in conferences to discuss attainment of literacy goals previously set and to formulate new ones for the future. Another is providing opportunities for students to interact with peers discussing texts generated by students or by others. The "best works" portfolios present a third opportunity. When students self-select these texts from their folders to include in their portfolios, they should be directed to reflect over past learning and to write an explanation of their selection process. On an ongoing basis, teachers can encourage students to maintain learning logs, response journals, and students' periodic written evaluations and reflections concerning their own literacy processes and products. In this way, students not only take responsibility for their learning but also develop the capacity to be reflective thinkers by making their literacy experiences the object of conscious thought (Vygotsky, 1962, 1978).

## Performance Assessment

Another purpose for assessment is to find out whether students have mastered literacy skills and are able to coordinate and apply them in new situations. This is the essence of performance assessment. Farr (1991) notes that assessment specialists have long urged

use of fewer paper-and-pencil tests and more performance measures. The combination of inertia, complexities of performance assessment development, and the several economies of paper-and-pencil instruments have made actual progress in performance assessment development in traditional educational settings minimal to date. (p. 479)

However, good teachers have long used performance tasks in their classrooms, but perhaps have not seen them as such a potent assessment vehicle. We ask students to take skills learned through literacy events, recombine them, and apply them in a new context. For example, upon completing a short story we might ask students to script it into a play using all the conventions of drama and standard English (as appropriate) and to produce and perform it as an assessment measure. Another example, of course, is the collection of writing found in the student's writing folder. For that matter, the New Jersey testing program already contains performance measures and is moving toward increased reliance on them. This movement is evident in many of the state testing programs, such as those in Michigan and Illinois. Thus, performance measures span the continuum from formal to informal assessment.



Performance assessments measure tasks that most closely mirror those required in the world beyond the classroom. Postsecondary educators and employers show little interest in isolated skills tests, even as an entry-level screening instrument. Instead, for example, secondary students write essays as part of their application to college. A magazine editor assessing a candidate for an editorial assistant position will require the person to edit a long passage for the conventions of English, to organize the disparate passages into a coherent and cohesive text, and to compose a text containing generative, well-organized ideas conveyed in standard English. Similar demands for integrating isolated skills into an orchestrated whole are evidenced in tasks requiring the automobile mechanic to read with comprehension the manufacturer's maintenance recommendations and communicate these to the consumer. All three performance tasks draw upon students' use of skills learned in school to develop the complex understandings that enable them to become literate citizens.

## Portfolios

Portfolios provide a fourth means to document and assess our students' literacy growth. As we examine the concept of portfolios and their contents, we recognize the interplay between the forms of assessment previously described and a fifth, testing, described below. Portfolios are defined as the collection of samples, artifacts, or documents that provide concrete evidence of students' literacy development during a specific period ranging from a semester to a year to several years. The collection contained in the portfolio might include teacher and student observations of the process leading to a performance that might also find its way into the portfolio. This artifact of the performance might be a piece of writing, a videotaped performance, or an audiotape of an interview. The concept of the portfolio for educational assessment is based on the model of the portfolio used by artists, photographers, and architects, who collect representative works that display their skills to the public. The work of Vera John-Steiner (1987), who collected samples of the verbal and visual thinking of more than a hundred creative thinkers, ranging from Darwin to Martha Graham, provides additional support for gathering representative samples that document students' literacy development.

Agencies from the National Assessment of Educational Progress (NAEP) to state testing departments as well as local school districts are looking with growing interest at student portfolios as a means to authentically document learning. The contents of a portfolio can provide much more compelling evidence than isolated grades and quantitative scores. Unfortunately, these agencies who have called for representative student portfolios have found that the concept of portfolios is often poorly understood. When the administrators of NAEP first requested samples of student portfolios in 1992, they found that short-answer worksheets and objective tests were all too often the major components of many student portfolios. Apparently, some teachers rely heavily on these kinds of materials to the exclusion of whole pieces of writing, extended responses to literature, and varied opportunities for speaking, listening, and meaningful viewing.

The contents of portfolios should reflect several literate behaviors by learners: (1) Students are able to collaborate with teachers and peers in selecting best and most representative works illustrating the student's growing proficiency in oral and written language as well as visual texts. (2) Students are able to arrange their works in ascending order of literacy proficiency. (3) Students are able to provide a journal, a diary, or a reflective essay explaining their rationale for the selection and order of their pieces. The portfolio can include a wide range of work: list of books read; photocopied pages from books read, such as favorite passages or periodic samples of materials demonstrating increasing complexity; samples of response to literature, including writing, illustrations, pages from response or dialogue journals; a strategies checklist; attitude or interest inventories; anecdotal records and conference notes; student self-evaluations; audio- or videotapes of students reading at different points in the year; photographs or tapes of projects; students' writing, such as poems, plays, personal narratives, responses to activities or events including movies, plays, trips, sports, and art; writing in the content areas; taped segments of students' oral language while engaged in small-group discussions, planning, and problem solving; and graphs, drawings, and other visual displays of information synthesized from literacy events.

Currently, districts and state agencies are wrestling with three questions: What criteria should be used to assess the portfolio? How long should a portfolio be maintained? Where should portfolios be housed for easy access? Possible criteria for assessment depend on the purpose. We may be looking for the breadth of literacy experiences represented in the portfolio, evidence of students' growing control of cognitive strategies while engaging in literacy events, or demonstrations of growing competencies in language fluency and style. Establishing the focus for portfolio assessment and the criteria to be used in assessing student learning must be grounded in clear standards. The work of such researchers as Tierney, Carter, and Desai (1991), Farr and Tone (1994), and Harp (1996) provide many models for establishing the criteria that can inform a district assessment instrument. However, it is not enough for a local district to identify the criteria and to create an instrument to guide portfolio assessment. The district must provide training so that teachers achieve a shared understanding of the criteria and uses of the instrument to assess the student growth evidenced in the portfolio. Districts must also provide opportunities for teachers to meet and discuss issues, such as the period that the portfolio will be used to document student growth (whether a year or all twelve years). New technologies permit long-term records through CDs and in other multimedia formats. However, some districts are not able to access these technologies yet, and for them, the physical logistics of housing the portfolio and moving it from teacher to teacher presents problems that must be solved.



## Hallmarks of Good Assessment

Meaningful measure of student progress is based on continuity of criteria across time and setting. When the criteria shift from one student to the next, or between test administrations, the data are not comparable and do not permit appropriate inferences about students' growth and development. Roland Case (1992) reminds assessors that the hallmarks of good assessment are also validity, fairness, triangulation, and enhancement of learning. Assessments are valid when they measure the performance they purport to measure and discriminate between different performance proficiencies. They are fair when students understand the standards by which they will be judged and have the opportunity to display proficiency in a variety of ways. Assessment results provide rich descriptions of student proficiency when they triangulate the data (Lincoln and Guba, 1985), that is, when they use several sources of information that document conclusions focused on specific, convergent aspects of students' literacy proficiencies. Finally, assessment in any form is indefensible if it does not contribute to student learning, whether directly or indirectly. Information from an assessment should inform planning and delivery of curriculum and instruction, but assessments should also afford students an opportunity to reflect on themselves as learners. If an assessment instrument does not offer a learning experience to the student, it serves merely as a gatekeeper.

## Tests

The preponderance of objective, norm-referenced tests traditionally have offered students little information about themselves as learners. However, the same could be said of the uninformed use of a teacher's pop quiz or the misuse of the portfolio as a mere paper repository. Traditional testing is akin to a behaviorist's view of the learner as the passive recipient of data. Current testing theory is based on the cognitive psychologists' view of the learner as an active construer of meaning from the information available from the environment. We now know, for example, that we should not try to decontextualize test items by using short excerpts in reading that block the reader's use of prior knowledge to construct new information. Short passages prevent skilled readers from using the reading strategies they would employ with a longer passage as they become familiar with the topic and discover the organization of the text. Current theory dictates the use of long passages across a variety of text types and topics to gain a valid indication of reader proficiency. We no longer depend solely on short answers, such as multiple choice, but include open-ended items that permit test takers more latitude to display their reading skills. Parallel issues arise in the assessment of writing. We no longer assume that students' abilities to revise and edit a given text reflect their abilities to generate, organize, and elaborate original ideas. In short, editing texts is not a complete test of writing proficiency. Current theory holds that any test that purports to be a valid test of writing must include opportunity for the writer to compose original, well-organized text with varied sentence structures and rich word choice using the conventions of standard written English.

New Jersey's new 4th-, 8th-, and 11th-grade tests, which are aligned to the language arts literacy standards, reflect much of current theory concerning learning and testing. Not only do they incorporate long reading passages with opportunities for open-ended responses to different text types and theme-based topics, but they also elicit multiple writing samples from students. In addition, they provide opportunities for students to integrate the reading and writing processes through decision making and problem solving in order to compose an original text using information from a reading passage as support. The tests also honor the hallmarks of assessment outlined by Case. They are valid because they measure what they purport to measure, that is, they provide rich contexts for the assessment of meaningful speaking, listening, writing, reading, and viewing behaviors. The new tests are also fair because they are aligned to the language arts literacy standards and indicators that have been published and distributed to educators, who will share them with their students, parents, and the community. Furthermore, this curriculum framework provides the same audiences with vignettes and activities that vividly translate the standards into classroom practices. Teachers can use this material to enhance student attainment of the standards and to foster student success on the new tests.

Educators are urged to share information from this document with parents to strengthen the home-school connection and involve parents in the development of their children's literacy. These steps have been taken to ensure the fairness of the New Jersey tests. Case identifies triangulation as the third hallmark of good assessment. Contemporary test theorists recognize that there are still limitations in a single assessment of a student's literacy development. Therefore, contemporary test makers, such as NAEP and the New Jersey statewide assessment program, are embracing performance assessments and portfolios as part of their procedures. Within the area of testing, these additional measures offer a means of triangulating the data by providing a second source of information about students' proficiency in literacy. The fourth requirement of good assessment speaks to assessment as an occasion for learning, an occasion

when a participant learns something about the nature of assessment. It is a moment when he or she suddenly...sees his or her work as someone else might...because the participant, himself or herself, steps outside and becomes an onlooker. (Wolf, 1993).

It is an occasion for engaging in what William James saw as a reflective process of turning away from our outward-looking point of view in order "to think of ourselves as thinkers" (in Wolf, 1993). The New Jersey tests' rich thematic approach allows students to read meaningful stories, practice higher-order thinking skills, and reflect upon and become conscious of their strategies as learners.

As teachers or district administrators, we should keep these four hallmarks in mind when we review commercial tests or construct our own. Moreover, we should be sure as we look at these tests (whether published or self-constructed) that they reflect current theories of learning and testing. We must also be sure that our tests are offering us information concerning the students' developing proficiencies in literacy and successful movement towards the standards and indicators. We must also remember that testing is only one of five major approaches to assessment. It cannot stand alone as the developers of NAEP and New Jersey's state assessment have found, but must be complemented by the panoply of assessment tools available to us and by adequate training in such techniques as observation, record keeping, effective use of student self-assessments, construction of challenging performance tasks, techniques for portfolio assessment, and test development and interpretation of test results. To become proficient assessors of their students, teachers may need systematic training that provides demonstrations, guided practice, and coaching in these techniques and in the development and use of rubrics for meaningful assessment. The vignettes presented in this document contain suggestions for *Possible Assessments*. A review of this material will reveal the many ways that teachers can assess students' literacy other than tests. Moreover, the sections called *Questions for Reflection* invite teachers to assess their own educational beliefs and practices as they deliberate over varied instructional approaches to the standards and indicators addressed in each vignette. Since skilled assessment is inextricably linked to good instruction and students' mastery of the standards and indicators, it is a critical piece of New Jersey's goal for New Jersey's learners. We must remember that we are assessing minute-by-minute and day-by-day and that these judgments impact upon learners, their views of themselves, and their literacy development.



## Assessment Strategies

STRATEGY	DESCRIPTION
<b>Observation</b>	The teacher observes students in a learning situation, looks for evidence of understanding and literacy development, and makes written notes about students' comments, insights, and behaviors. Anecdotal records allow teachers to record the rich detail available in ongoing and cumulative observations of literacy processes and products in a flexible, open-ended manner.
<b>Written records</b>	Teachers, like students, keep written records of assessment data collected from the range of methods. These records may be made on removable notes during observation or more sustained discussions of literacy development. Teachers can create a folder for each student or use a binder or spiral notebook with three to four pages allocated to each student in which she can write comments, store checklists and surveys, and/or attach the removable notes with her observations.
<b>Questioning</b>	<p>Questions are asked to evaluate students' thinking and reasoning. Open-ended questions targeting higher-order thinking skills require students to think about literacy events and provide opportunities to discover and validate ideas.</p> <p>Examples:            What would happen if...?            How do you know that...?            Can you predict...?            Why do you think...?</p>
<b>Checklists</b>	Checklists provide a succinct, easy-to-record method for documenting and/or recording student behaviors in reading and writing.
<b>Informal reading inventories</b>	Reading inventories allow teachers to sample the oral reading of their students for diagnostic purposes and to plan future instruction. The informal reading inventory and the reading miscue analysis are two instruments available to teachers for this purpose.
<b>Interview guides</b>	Interview guides are published or locally designed (by students or teachers) instruments that serve a variety of purposes in the class. They may be used for interviewing classmates, the teacher, or guest speakers to provide students practice in clear, oral speaking and listening. Interviews may elicit information about the interviewee's background, knowledge, interests, and attitudes about a topic.

**Informal protocols**

Informal but informed one-on-one assessment reveals the strategies students use to construct meaning as they speak, listen, write, read, and view. Measures such as reading miscue analysis or inspection of students' writing for a pattern of development are two examples.

**Surveys**

Surveys provide direct questions or rating scales that elicit information about teachers' or students' processes, products, attitudes, and interests.

**Conferences**

Teacher and student or peer groups discuss the student's literacy experiences. Discussion will focus on selection of goals, topics, forms, strategies, and assessments. Conferences can be formal (scheduled) or informal (impromptu).

**Performance-based tasks**

These are multiple-exercise problems, presented in real-life situations, that require students to apply previously learned skills. Performance-based tasks are completed by individuals or by groups of students.

**Self-assessments**

Self-assessment is a process that allows students to participate in setting goals and monitoring learning. Teachers encourage students to maintain learning logs, response journals, and students' periodic written evaluations and reflections concerning their own literacy processes and products. This kind of exercise invites students' metacognitive awareness by making students' literacy experiences the object of conscious thought.

**Tests**

Tests are structured instruments designed to measure what students know at a specific period in development. They may be norm-referenced or criterion-referenced, published or locally designed, voluntary or mandated samples of literacy skills. Tests may elicit objective, open-ended, or essay responses. Students may participate in the construction of locally designed tests.

**Literacy folders**

Literacy folders contain daily work samples from the students' literacy experiences with verbal (oral or written) and visual texts. These samples may be drafts or finished products. The folder serves as the source for work selected for the portfolio.

**Portfolios**

Portfolios are defined as collections of samples, artifacts, or documents that provide concrete evidence of students' literacy development during a specific period ranging from a semester to a year or even several years. The contents are selected by the student, sometimes in collaboration with the teacher, from the literacy folder as the student's best work or according to some other criterion agreed on by the student and the teacher. The portfolio also contains a reflective explanation by the student offering the rationale for the selection and order of the pieces.

## References

- Case, R. (1992, Winter).  
On the need to assess authentically.  
*Holistic Education Review*, 5(4), 14–23. Garden City, NY: Adelphi University.
- Farr, R. (1991).  
Formal methods of evaluation. In J. Flood, J. Jensen, D. Lapp, & J. Squire (Eds.),  
*Handbook of research on teaching the English language arts*. New York: Macmillan.
- Farr, R., & Tone, B. (1994).  
*Portfolio performance assessment: Helping students evaluate their progress as readers and writers*.  
Fort Worth, TX: Harcourt Brace College.
- Goodman, Y. (1991).  
Informal methods of evaluation. In J. Flood, J. Jensen, D. Lapp, & J. Squire (Eds.),  
*Handbook of research on teaching the English language arts*. New York: Macmillan.
- Harp, B. (1996).  
*The handbook of literacy assessment and evaluation*. Norwood, MA: Christopher-Gordon.
- Hiebert, E. H. (1991).  
Teacher-based assessment of literacy learning. In J. Flood, J. Jensen, D. Lapp, & J. Squire (Eds.),  
*Handbook of research on teaching the English language arts*. New York: Macmillan.
- James, W. (1890, 1950).  
*The principles of psychology*. New York: Dover.
- John-Steiner, V. (1987).  
*Notebooks of the mind: Explorations of thinking*. Albuquerque, NM: University of New Mexico Press.
- Lincoln, Y. S., & Guba, E. G. (1985).  
*Naturalistic inquiry*. Newbury Park, CA: Sage.
- Tierney, R. J., Carter, M. A., & Desai, L. E. (1991).  
*Portfolio assessment in the reading-writing classroom*. Norwood, MA: Christopher-Gordon.
- Vygotsky, L. (1962).  
*Thought and language, language and thought*. Cambridge, MA: The MIT Press.
- Vygotsky, L. (1978).  
*Mind in society—The development of higher psychological processes*.  
Cambridge, MA: Harvard University Press.
- Wolf, D. P. (1993).  
Assessment as a mode of learning. In R.E. Bennett & W. C. Ward (Eds.),  
*Construction vs. choice in cognitive measurement*. New Jersey: Lawrence Erlbaum.

